

Analysis of major failures in Europe's power grid

Martí Rosas-Casals^{*,a,b}, Ricard Solé^{b,c}

^a*Càtedra UNESCO de Sostenibilitat, Universitat Politècnica de Catalunya (UPC),
EUETIT-Campus Terrassa, Edif. TR4, C. Colom, 1, 08222 Barcelona, Spain*

^b*ICREA-Complex Systems Lab, Universitat Pompeu Fabra - PRBB, Dr. Aiguader 88,
08003 Barcelona, Spain*

^c*Santa Fe Institute, 1399 Hyde Park Road, New Mexico 87501, USA*

Abstract

Power grids are prone to failure. Time series of reliability measures such as total power loss or energy not supplied can give significant account of the underlying dynamical behavior of these systems, specially when the resulting probability distributions present remarkable features such as an algebraic tail, usually considered the footprint of self-organization and the existence of critical points. In this paper, seven years (from 2002 to 2008) of Europe's transport of electricity network failure events have been analyzed and the best fit for this empirical data probability distribution is presented. With the actual span of available data and although there exists a moderate support for the power law model, the relatively small amount of events contained in the function's tail suggests that causal factors other than self-organization or a critical state might be significantly ruling these systems' dynamics.

Key words: power grid, complex networks, time series

PACS: 84.70.+p, 89.75.Hc, 95.75.Wx

2000 MSC: 62-07

1. Introduction

There has been in recent years an increasing awareness about infrastructure networks security and reliability [1, 2, 3]. Modern society's functional capacity relies on an optimal operation of infrastructure and information

*Corresponding author

Email addresses: `rosas@mmt.upc.edu` (Martí Rosas-Casals), `ricard.sole@upf.edu` (Ricard Solé)

networks such as roads, railways, gas and oil pipes or Internet. Particularly vital, and at the same time prone to failure, are electric power transmission networks. These are extremely complex engineered systems, composed of multiple and interconnected elements, whose reliability depends both on each component's behavior and, at the same time, on the many different dynamical interactions that span over and rule the overall connectivity of the system.

Although it is not always the case, a malfunction of a power transmission system shows usually itself as a blackout. This is a direct consequence of a cascading failure involving several of its composing and linking elements. This fact turns the study of the details of failures in power transmission networks from a traditional engineering point of view a hard task, if not an impossible one most of the times. In order to reduce the inherent complexity of this detailed approach, some new ways have been proposed in recent years. One of them is that of ignoring the details of particular failures and to focus on the study of global behaviors and dynamics of time series with approximate global models. Concepts such as criticality and self-organization have been applied to characterize blackout data, suggesting that the frequency of large blackouts is governed by non trivial distribution functions such as power laws and, consequently, that power systems are designed and operated near a critical point. (For a comprehensive review on this approach, see Ref.[4] and references therein).

This paper analyses for the first time, and as far as we know, the statistics of major electric transmission network events in the European power grid from this aforementioned complex systems approach. Following essentially the statistical analysis presented in Ref.[5], we estimate the basic parameters of the power-law model, then calculate the goodness-of-fit between the data and the power law and finally we compare the power law with alternative hypotheses via a likelihood ratio test. The paper is organized as follows. In section II European major events data is presented and explained. In section III blackout data is analyzed. Finally, section IV summarizes our main results.

2. UCTE major events data

European power network reliability data can be found in the Union for the Co-ordination of Transmission of Electricity (UCTE) web page, publicly available from 2002 onwards in monthly statistics format [6]. The UCTE

is the association of Transmission System Operators (TSOs) in continental Europe and manages data from 24 different European countries. Due to the complexity of events, sometimes involving more than one TSO, types of interruptions in the network and short time given to gather this information, UCTE major events data is somehow limited in its scope and does not provide a fully detailed account of some events. It is, nonetheless, the best documented source that has been found. For each major event, it summarizes the following information:

- **Country.**
- **Substations involved.**
- **Reason (R).** Broadly classified into four groups: (1) overloads (also calculated brakes), (2) failures (false operation, failure in protection device or other element), (3) external (outside impacts and very exceptional weather and natural conditions) and (4) other or unknown reasons.
- **Energy Not Supplied (ENS).** Measured in MWh, as loss of energy from the consumption side.
- **Total Loss of Power (TLP).** Measured in MW, as loss of production from the generation side.
- **Restoration Time (RT).** Measured in minutes. Note that since ENS and TLP are measured from different sides, RT can not be assumed as the ratio of ENS over TLP. It can be considered, therefore, an independent reliability measure.
- **Equivalent Time of Interruption (ETI).** Defined as the duration of an interruption in minutes multiplied by the energy not supplied divided by the consumption for the last 12 months. Defined in this way, the ETI allows a direct comparison between TSOs in terms of interruptions that occurred during a year.

From 2002 to 2008, both years inclusive, 908 major events have been noticed. Due to the complexity of events some entries have zero value in one or more of their categories. While these zeroed values have not been considered, the rest of numerical values are effective measures of major events occurred in the UCTE power grid and, consequently, they have all been included in order to develop the analysis presented in this paper.

3. Probability distribution analysis

The study of the statistics and dynamics of series of events with approximate global models has been one of the most popular topics in the last twenty years, specially within the interdisciplinary study of complex systems. Probability distribution functions with a heavy tailed dependence in terms of event or object sizes seem to be ubiquitous statistical features of self-organized natural and social complex systems [7]. The appearance of algebraic distributions, specially power laws, is often thought to be the signature of hierarchy, robustness, criticality and universal underlying mechanisms [8]. Electric power transmission networks have not escaped this captivation for power laws quest. Time series of usual measures of blackout size like energy unserved, power loss or number of customers affected, have been shown to be algebraically distributed in North America [9], Sweden [10], Norway [11], New Zealand [12] and China [13]. This apparent ubiquitous evidence have led to believe and try to demonstrate that power systems (a) tend to self-organize near a critical point and (b) that there may be some universality ruling the inner depths of these systems.

In spite of this *evidence*, most of the aforementioned literature rely on poorly performed statistical analysis and results can not be trusted. In some cases methodologies are not clearly explained (i.e., Ref.[11]) or simple visual inspection can clearly dismiss the analysis performed to rule in the power law hypothesis (i.e., Ref.[10] and Ref.[13]). In other cases, proper usage of statistical tools have given new results that limit the scope of the original analysis (i.e., Ref.[9] is dismissed as insufficiently substantiated in Ref.[16] and reanalyzed in Ref.[5], finding moderate support for the power law hypothesis and even some for an exponential distribution).

In this section we analyze the probability distributions of three malfunction measures of the European power grid: energy not supplied (ENS), total loss of power (TLP) and restoration time (RT). The results are summarized in Table 1 and shown in Figure 1. The methodology that has been used is that described in Ref.[5], where a maximum likelihood approach is proposed to estimate the heavy tailed function from the data and a significance test is constructed for testing the plausibility of the power law. Measures shown in Table 1 are generic statistics on one side and results of the aforementioned statistical analysis on the other. The analysis of the power law distributed data has been performed in the following way. We assume a quantity x follows a (continuous) power law distribution if it is drawn from a probability

Data set	n	$\langle x \rangle$	σ	x_{max}	Maximum likelihood				Support for PL
					\hat{x}_{min}	$\hat{\alpha}$	n_{tail}	p	
ENS	690	552	7004	180000	185 ± 72	1.7 ± 0.1	104 ± 120	0.24	Moderate
TLP	576	400	1790	26746	615 ± 244	2.1 ± 0.2	57 ± 96	0.36	Moderate
RT	897	510	3328	44640	150 ± 68	1.69 ± 0.07	157 ± 115	0.73	Ok

Table 1: UCTE major events generic statistics and power law fits. For each measure we give the number of occurrences n , mean $\langle x \rangle$, standard deviation σ , maximum observed occurrence x_{max} , lower bound to the power law behavior \hat{x}_{min} , scaling parameter value $\hat{\alpha}$, occurrences in the power law tail n_{tail} and p value p . The last column indicates the support for whether the observed data is well approximated by a power-law distribution. Estimated uncertainties for \hat{x}_{min} , $\hat{\alpha}$ and n_{tail} are also shown.

density $p(x)$ such that $p(x) \propto x^{-\alpha}$, where α is the scaling parameter of the distribution. Since the probability density of a power law distribution diverges as $x \rightarrow 0$, there must exist a lower bound to the power law behavior [16]. We denote this lower bound as x_{min} and the number of events contained in the upper range as n_{tail} .

In order to estimate both parameters α and x_{min} , we first use the maximum likelihood estimator for the scaling parameter for the continuous approximation, which is defined as

$$\hat{\alpha} = 1 + n \left[\sum_{i=1}^n \ln \frac{x_i}{x_{min}} \right]^{-1} \quad (1)$$

where x_i , $i = 1, \dots, n$, are the observed values of x such that $x_i \geq x_{min}$. The former estimator depends on x_{min} and its value is not known a priori. Before calculating $\hat{\alpha}$, we need to first discard all samples below this point so that we are left with only those for which the power-law model is valid. The Kolmogorov-Smirnov (KS) statistic is used in this sense, which is defined as the maximum distance D between the cumulative distribution functions of the data $S(x)$ and the fitted model $P(x)$:

$$D = \max_{x \geq x_{min}} |S(x) - P(x)| \quad (2)$$

Our estimate \hat{x}_{min} is the value of x_{min} that minimizes D .

Given an observed data set and a hypothesized power-law distribution from which the data are drawn, one second step is knowing whether the power

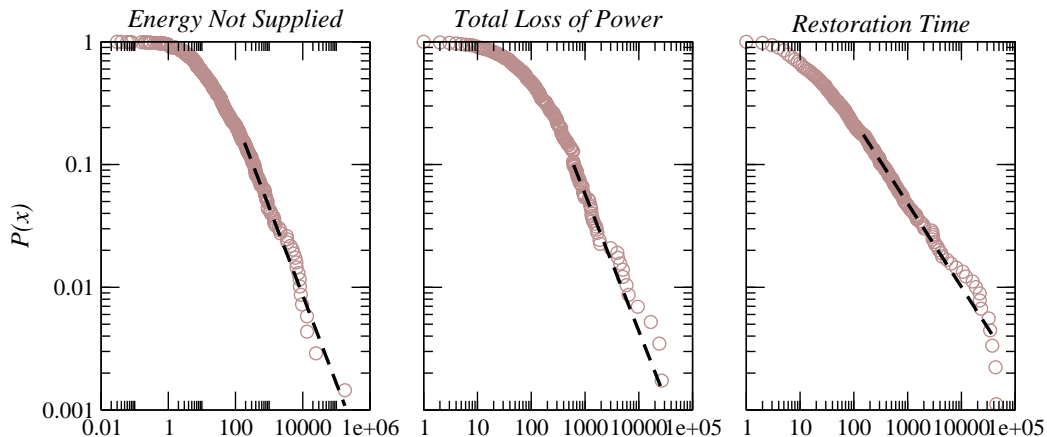


Figure 1: Cumulative distribution functions $P(x)$ and their maximum likelihood power law fits (discontinuous straight line) for the UCTE reliability measures energy not supplied, total loss of power and restoration time.

law hypothesis is a plausible one. The aforementioned KS statistic is used again as a goodness-of-fit test between real data and synthetically generated power-law distributed data. Each synthetic data set is fitted individually to its own power-law model and the KS statistic for each one relative to its own model is calculated. The fraction of the time that the resulting statistic is larger than the value for the empirical data is the p -value that we are looking for. It denotes the significance test result and quantifies the plausibility of the hypothesis. The power law is ruled out if $p \leq 0.1$.

As we can see in Table 1, the power law model is a plausible one for every data set considered (i.e., the p -value for the best fit is sufficiently large) and the scaling parameter values are similar to those encountered in the literature for the ENS and TLP distributions [4, 14]. Yet the power law model explains only a small amount of events: 15% for ENS ($n_{tail} = 104$), less than 10% for TLP ($n_{tail} = 57$) and 17% for RT ($n_{tail} = 157$) (even though it holds the better fit, with $p = 0.73$). We believe that measures such as n_{tail} and x_{min} are fundamental to estimate the span of the power law behavior and to develop further quantitative models, yet these values have not been considered in any of the aforementioned references. Only in the reanalysis of Ref.[9] done in Ref.[5] we have found an estimate for n_{tail} that gives an explanation for a 28% of the events.

We assume that the limited span of available data in each set might have a sensible influence in the final power law fitting outcome. It is nonetheless

Data set	log-normal		exponential		Weibull		PL + cut off	
	<i>LR</i>	<i>p</i>	<i>LR</i>	<i>p</i>	<i>LR</i>	<i>p</i>	<i>LR</i>	<i>p</i>
ENS	-0.405	0.68	2.64	0.00	-0.393	0.69	-0.419	0.36
TLP	0.319	0.75	3.42	0.00	0.467	0.64	-0.08	0.68
RT	-0.382	0.70	7.57	0.00	-0.329	0.74	-1.7	0.06

Table 2: Test of power-law behavior. Positive values of the likelihood ratios *LR* favors the power law model. Values of $p \geq 0.1$ imply though that result can not be trusted. The exponential distribution is definitely ruled out as possible model and only for the restoration time *RT*, the power law with cut off could be considered a valid model.

evident from these results that (a) power law behavior can not be assumed for the whole data observed, (b) we can not accept the existence of any critical point at this stage of the data span and (c) there must be right now considerably more dynamics not explained by the power law model.

Following Ref.[5], as a third and final step, and in order to check if other distributions may be a better fit, a likelihood ratio (LR) test has been performed. The essential idea behind the LR test is to compute the likelihood of the data under two competing distributions. The one with the higher likelihood is then the better fit. Since the sign of the LR alone do not definitively indicate which model is the better fit (because it is subject to statistical fluctuation), we estimate the size of the expected fluctuations (i.e., standard deviation) using Vuong’s method [17]. This method gives a new *p*-value that tells us whether the observed sign of the LR is statistically significant. Likelihood ratios and *p*-values results with respect to log-normal, exponential, stretched exponential (Weibull) and power law with cut off distributions are shown in Table 2. Positive likelihood values favor the power law hypothesis and *p*-values higher than 0.1 imply no significance on the results. Log-normal, Weibull and exponential distributions can be ruled out since $p \geq 0.1$ in the first two and $p = 0$ in the latter with positive LR values. Power law with cut off can be ruled out in the ENS and TLP data sets though it is a plausible option for the RT data set.

4. Summary and discussion

Power outages are considered unexpected phenomena in power grids. They appear without warning and, though widely investigated, there is not a

common accepted theory that explains neither their pervasiveness nor their inner dynamics. The statistical overabundance of big blackouts has been explained using theories of systems failure able to reproduce their empirically found probability distribution. This distribution is considered a power law for most of the literature encountered, with the consequences that this algebraic tail involves (i.e., self-organization, criticality and universalities). In order to add one more reference to this field, in this paper we have analyzed seven years of disturbances data for the UCTE power grid and for three major event measures: energy not supplied, total loss of power and restoration time. Although evidences for self-organized criticality have been suggested for even such as a short time span as five years of data [15], the results presented in this paper support the power law hypothesis moderately and only for two (ENS and TLP) of the three measures considered. Moreover, the amount of events explained by the power law hypothesis can be considered negligible. These facts make it difficult to accept the existence of an equilibrium point near criticality for the UCTE power grid, at least at this stage of data analysis, and it also suggests that most of the power grid dynamics should be explained by different dynamical models, other than those that generate power laws as their statistical outcome.

There still exist many complexities not explained in this systems. Ongoing research is now focused in analyzing major events probability distributions in connection with (a) the reasons that trigger these major events and (b) the structure and topology of the power grids involved in these major malfunctions. Results will be published elsewhere.

We deeply thank David Chury, Data and Modeling Expert at the UCTE, for his patience, kindness and useful discussions. We also thank Aaron Clauset for enlightening comments and clarifications about distributions with heavy tails. One author (MRC) wishes to thank specially Selina Delgado-Raack for most helpful conversations and support. This work has been supported by the Santa Fe Institute and the Technical University of Catalonia, UPC.

References

- [1] Wong, A. & Yeung, A., *Network Infrastructure Security*, Springer Science, (2009).
- [2] Lewis, T. G., *Critical Infrastructure Protection in Homeland Security*, John Wiley & Sons, Inc., (2006).

- [3] Perrow, C., *The Next Catastrophe*, Princeton University Press, (2007).
- [4] Dobson, I., Carreras, B. A., Lynch, V. E., & Newman, D. E., *Chaos*, **17**, 2, (2007).
- [5] Clauset, A., Shalizi, C. R. & Newman, M. E. J., *SIAM Review*, **51**, 4, (2009).
- [6] <http://www.ucte.org/resources/publications/monthlystats>
- [7] Sornette, D., *Probability Distributions in Complex Systems*, arXiv:0707.2194, (2007).
- [8] Bak, P., *How nature works : the science of self-organized criticality*, Copernicus, (1996).
- [9] Carreras, B. A., Newman, D. E., Dobson, I., & Poole, A. B., *IEEE Transactions on Circuits and Systems I*, **51**, (2002).
- [10] Holmgren, A. J., & Molin, S. *Journal of Infrastructure Systems*, **12**, 4, (2006).
- [11] Bakke, J. O. H., & Kertész, J., *Europhysics Letters*, **76**, 4, (2006).
- [12] Jordan, C., Rasmussen, H. & Suri, R., <http://www.maths-in-industry.org/miis/76>, Study Group Report, (2006).
- [13] Weng, X., Hong, Y, Xue, A. & Mei, S., *Journal of Control Theory and Applications*, **4**, 3, (2006).
- [14] Chen, J., Thorp, J. S., & Parashar, M., *Proceedings of the 34th Hawaii International Conference on System Science*, (2001).
- [15] Carreras, B. A., Newman, D. E., Dobson, I., & Poole, A. B., *Proceedings of Hawaii International Conference on System Science*, (2000).
- [16] Newman, M. E. J., *Contemporary Physics*, **46**, (2005).
- [17] Vuong, Q. H., *Econometrica*, **57**, 2, (1989).